ATTACHMENT "A"

Report to AD/RR on Feasibility Study

of

Minicard for Mass Data Problems

1. Purpose of the Test:

   a. Specific:

      (1) To measure by actual test what proportion of a mass data storage, recall, and manipulation problem can be done by existing Minicard facilities.

   b. Side Benefits Expected for ORR:

      (1) To gain general experience in code indexing methods, and also learn the unique problems of preparing material for input into the Minicard system.

      (2) To obtain the basis for a fairly reliable estimate of the manpower requirements necessary for input of the entire mass data base, should the Minicard system prove satisfactory for the purpose.

   c. General Benefits:

      (1) In addition to the above specific measures of Minicard capabilities, we expected the test to provide us with experience which would be valuable in connection with processing data for any mass data storage and manipulation equipment, including the latest computers.

2. Background and Character of Data:

   a. [redacted]

25X1D

NSA

25X1D1a 25X1D1a 25X1D1a

25X1D1a
& NSA

[redacted] The time span of the data sample covered the period January - mid-February 1958.

3. General Approach:

a. Since we needed to have the data output in ordered categories, in order to search for suspected interrelationships, our first step was to arrange for the data to be stored by fixed field within each Minicard. Without such fixed fields, formatted output would not have been possible. Moreover, we would not have been able to put our data in order without fixed fields. Next, we designed a transcript sheet for the purpose of transferring the information in ordered format from the reports, so that data, in turn, could be converted into machine language for entry into the system.

4. The Input:

a. During the first week the information was transcribed on the transcript sheet, the cards were punched manually from the sheets, and these punched cards were converted mechanically to paper tape. It quickly became apparent that the [redacted] 25X1D
25X1D [redacted] data had such readable continuity that the paper tape NSA
NSA could be punched at the same time that the information was being analyzed. By this short-cut, three steps were saved in NSA
the process of input: (1) transcribing the information to transcript sheets; (2) punching and, (3) verifying the cards NSA
from which the paper tape was made. By the shortened method, information from slightly less than [redacted] was NSA 25X1D
punched into paper tape at about [redacted] com- NSA 25X1D
25X1D pared to about [redacted] before. After the [redacted] had been NSA 25X1D
punched into the paper tape, it was fed into the Minicard
NSA camera where the punched hole information was read from the tape and photographed on Minicard film in binary coded form.
Next, the Minicard file reels containing all the [redacted] in- 25X1D
formation were put through the film processor, then to the NSA
waxer and then to the cutter which made them into individual
Minicards. Each Minicard contained information on one [redacted] 25X1D
25X1D [redacted] The Minicard test file was now ready for interroga-
NSA tion via the selector. NSA

5. Specific Steps for Testing Minicard Capabilities:

For this initial test the I/GM analyst, [redacted] formulated nine "questions" which I/GM analysts felt would test various aspects of the Minicard system. These questions and the results follow: 25X1A9a

a. Select from the file the following [redacted] NSA 25X1D
25X1D1a [redacted] List the selected NSA 25X1D
numbers by date (minor sort) and then by [redacted] (major sort) and print-out. NSA 25X1D

Results: 74 Minicards were selected out of the total file.

b. Select the following cities: [redacted] 25X1D1a
25X1D1a [redacted]
Sort the selected cards by date.

Results: No hits.

c. Select the following cities: [redacted] as a constant, 25X1D1a
25X1D1a [redacted] as variables, sort the selected cards by date.

Results: No hits.

d. Select the following cities: [redacted]
with both cities as constants, sort by date.

Results: No hits.

e. Select the following cities: [redacted] with 25X1D1a
both cities as constants, sort by date.

Results: No hits.

f. Select the following cities: [redacted]
25X1D1a [redacted]
Sort by date.

Results: No hits.

g. Select the following cities: [redacted] as a constant, 25X1D1a
25X1D1a [redacted] as variables, sort by date.

Results: No hits.

25X1D1a

h. Select: [redacted] no variables, sort by date.

Results: 83 hits.

25X1D1a i. Select: [redacted], no variables, sort by date.

Results: 133 hits.

6. Interrogation Results:

a. Out of nine questions asked, five groups of hits were registered. Four of the selected groups were sorted and listed by date. The fifth selected group registered only one hig, which was not worth printing out. The absence of hits on four questions indicates that no [redacted] activity pertinent to these questions was recorded in the published material used during the short date-time span involved.

25X1D

NSA

7. Strengths of the Minicard System:

a. The feeding of the paper tape into the camera for photographing and the following steps of developing the film and having it waxed and cut were performed at a fairly rapid rate. About $\frac{1}{2}$ minute was required to photograph one Minicard and 45 minutes to develop one reel of 200 feet containing 1,4000 Minicards. Another 30 minutes were required for waxing, and 30 more minutes for viewing the reel. At this juncture, the Minicards were ready for interrogation via the selector. The selector has a memory capacity of 17-19, six-character words which means that 17-19 words can be used in any combination for selection purposes such as constants, variables, etc. In this test, the memory capacity of the selector was more than adequate to interrogate the test file of Minicards, but this vocabulary does severely limit the complexity of initial questions that would probably be asked in mass data manipulation. The sorter, at peak, is relatively fast compared to standard IBM equipment, passing Minicards at a speed of 750 per minute. Both selection and sorting speeds were adequate for the test sample, but we believe both would be unsatisfactory for handling mass data of any magnitude. The Minicard size, of course, lends itself to economical storage of individual records. Only about one third of the Minicard was used to store data [redacted] which leaves a considerable potential of added information per Minicard.

25X1D

NSA

8. Weaknesses of the Minicard System:

a. The main weaknesses of the Minicard System were: (1) mechanical failures of individual pieces of equipment; (2) generally moderate speed of individual steps in data handling; and (3) slow and unformatted print-out. The periodic breakdown of the individual pieces of equipment due to some mechanical failure, repeatedly delayed the processing steps of this test. For example, when some reels of film had been developed as a result of input to the camera, it was found that the camera had not been functioning correctly, which necessitated a reshooting of the information. In the sorting and selection process, on more than one occasion both of these machines performed erroneously, which required a re-run of the same information.

b. Print-out via the duplicator is done by a standard Flexowriter, which gives rise to additional shortcomings. First, the Flexowriter prints at a very slow rate. For example, on one of the selections where 133 hits were made from the sample, these selected Minicards required one hour and ten minutes print-out time through the Flexowriter. This is a very unsatisfactory print-out rate for a mass data problem. For example, if all of the slightly less than [redacted] movement used in the test, which represent less than 1% of the total potential data volume, were to be printed out, it would take about 25 hours. 25X1D NSA

c. The other deficiency of the Flexowriter is that the information cannot be printed out in an ordered format. Each Minicard prints out across the paper and into the second line if necessary. Thus, the information from the second line may or may not be related to the information in the same position on the first line. A user of the listing looking down a given column could often find a jumble of information in the same column. This could be a great handicap in using this form of information listing, especially when large quantities of data are involved. We could, of course, have limited the number of data columns to the capacity of the Flexowriter carriage. Or, we might have used 2 Minicards for the data which we put on one. Either of these methods would have obviated spillover into a second line. However, the first method would have been very unsatisfactory from the analytic point of view, and the second method would have required examining two Flexowriter sheets side by side. Both methods would have been feasible for this test, but neither would have been satisfactory for handling a mass data problem of any magnitude.

9. Primary Conclusions:

In light of our test experience, we believe that the existing Minicard system is not an effective or an efficient means for storing, retrieving, and manipulating for sub-analytic purposes a data mass of any magnitude.

10. Bonus Benefits:

The experience gained in the code indexing methods of this selected portion of the mass data base improved considerably our understanding of the problems inherent in the preparation of such material for input into the Minicard system. In addition, the thirty days' experience demonstrated that a formidable investment of manpower will be required in any systematic effort on the part of ORR to process the entire [redacted] data base into any data system.

25X1D1a & NSA

Our understanding of the kinds of problems involved in any application of ADP equipment and procedures to selected ORR analytic problems was considerably increased.

The I/GM analysts improved their knowledge regarding both the specific intelligence potential of the [redacted] data, and the use of selective manipulation as a means of exploiting mass data.

25X1D1a & NSA